EAST Search History

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	3252	(715/513).CCLS.	US-PGPUB; USPAT; JPO; DERWENT	OR	OFF	2007/05/31 15:08
L2	0	("L1and(tag$1NEAR3extract$3)").PN.	US-PGPUB; USPAT; JPO; DERWENT	OR	OFF	2007/05/31 15:09
L3	130	L1 and (tag$1 NEAR3 extract$3)	US-PGPUB; USPAT; JPO; DERWENT	OR	ON	2007/05/31 15:11
L4	78	L3 and (@ad<"20010803" or @rlad<"20010803")	US-PGPUB; USPAT; JPO; DERWENT	OR	ON	2007/05/31 15:13
L5	18	L4 and (font and size and bold)	US-PGPUB; USPAT; JPO; DERWENT	OR	ON	2007/05/31 15:12
L6	2669	tag$1 NEAR3 extract$3	US-PGPUB; USPAT; JPO; DERWENT	OR	ON	2007/05/31 15:12
L7	80	L6 and (font and size and bold)	US-PGPUB; USPAT; JPO; DERWENT	OR	ON	2007/05/31 15:12
L8	48	L7 and (range$1)	US-PGPUB; USPAT; JPO; DERWENT	OR	ON	2007/05/31 15:12
L9	33	L8 and (@ad<"20010803" or @rlad<"20010803")	US-PGPUB; USPAT; JPO; DERWENT	OR	ON	2007/05/31 15:13

5/31/07

PORTAL
USPTO

THE ACM DIGITAL LIBRARY

Feedback Report a problem Satisfaction survey

Terms used **tag extraction HTML**   Found **11,881** of **201,890**

Sort results by   [relevance ▼]   📗 Save results to a Binder   Try an Advanced Search
                                       🔲 Search Tips   Try this search in The ACM Guide
Display results   [expanded form ▼]   ☐ Open results in a new window

Results 1 - 20 of 200   Result page: **1** 2 3 4 5 6 7 8 9 10   next
Best 200 shown   Relevance scale ☐ ☐ ◪ ◼ ◼

1 Data extraction: Fully automatic wrapper generation for search engines   ◼
       Hongkun Zhao, Weiyi Meng, Zonghuan Wu, Vijay Raghavan, Clement Yu
       May 2005 **Proceedings of the 14th international conference on World Wide Web
                 WWW '05**
       Publisher: ACM Press

       Full text available: 🔲 pdf(315.59 KB)   Additional Information: full citation, abstract, references, citings, index
                                                                         terms

       When a query is submitted to a search engine, the search engine returns a dynamically
       generated result page containing the result records, each of which usually consists of a
       link to and/or snippet of a retrieved Web page. In addition, such a result page often also
       contains information irrelevant to the query, such as information related to the hosting
       site of the search engine and advertisements. In this paper, we present a technique for
       automatically producing wrappers that can be used to extr ...

       Keywords: information extraction, search engine, wrapper generation

2 Data extraction: Web data extraction based on partial tree alignment   ◼
       Yanhong Zhai, Bing Liu
       May 2005 **Proceedings of the 14th international conference on World Wide Web
                 WWW '05**
       Publisher: ACM Press

       Full text available: 🔲 pdf(357.50 KB)   Additional Information: full citation, abstract, references, citings, index
                                                                         terms

       This paper studies the problem of extracting data from a Web page that contains several
       structured data records. The objective is to segment these data records, extract data
       items/fields from them and put the data in a database table. This problem has been
       studied by several researchers. However, existing methods still have some serious
       limitations. The first class of methods is based on machine learning, which requires
       human labeling of many examples from each Web site that one is interested in ...

       Keywords: data extraction, data record extraction, wrapper

3 Automatic information extraction from large websites   ◼
       Valter Crescenzi, Giansalvatore Mecca
       September 2004 **Journal of the ACM (JACM)**, Volume 51 Issue 5